

基礎研究は「時代」とともに在り

基礎研究の中から生み出される新発見、新発明によってイノベーションが起こり、産業構造と生活様式の変革につながることがあります。しかし、そうした成功事例はまれであり、基礎研究には高いリスクが伴うのも事実です。本稿では、NTTコミュニケーション科学基礎研究所の活動から生まれた技術を歴史の変遷とともに分析することによって、将来に向けた基礎研究の推進戦略と、取り組むべき課題を過去の成功事例から明らかにします。

まえだ えいさく

前田 英作

NTTコミュニケーション科学基礎研究所 所長

NTTコミュニケーション科学基礎研究所（CS研）が設立されて23年、そしてNTT再編に伴う新体制下で基礎研究を始めて14年が経ちました。この間にCS研の確たる骨格が形成され、基礎研究から生まれた技術が少しずつ世の中で使われるようになってきました。そして私たちは今、次の一步を踏み出すための大きな転換期にあると考えています。本特集では、「CS研オープンハウス2014」で展示した研究成果の中から7件を選び、技術の詳細と今後の展開について紹介しています。本稿ではそれらに先立って、研究開発における基礎研究の特質と役割について論じるとともに、イノベーションにつながる基礎研究のあり方とその推進戦略について触れたいと思います。

情報技術の「見えない」イノベーション

技術革新そのものがイノベーションであった蒸気機関や活版印刷機などの機械技術と比べると、情報技術は少し特殊な位置にあります。ある製品やサービスを指して、イノベーションが起きたと世の人はいうけれど、その中に情報技術も使われているはずだけれど、情報技術の重要性が世の中の人に

伝わっていないと、情報の専門家は嘆きます。それが、「情報技術の見えない」イノベーションです。また、Googleの検索エンジンやAppleのiPhoneを日常的に利用していたとしても、どんな情報技術がそうした製品やサービスの根幹をなしているのかを的確に語ることは必ずしも容易ではありません。ノーベル賞の対象となったiPS細胞（人工多能性幹細胞）や、音楽のあり方を変えたともいわれる「初音ミク」の裏側で情報技術がどのような役割を果たしてきたのか、一般の人は知らないのではないのでしょうか。新しい科学的発見や日々生まれる新しい製品、サービスが私たちの生活に大きな影響を与えイノベーションが起きているにもかかわらず、それに貢献している情報技術が表には見えていないのです。

一方、製品やサービスに直接、あるいはすぐには反映されないところで、情報技術の各分野で何年かに一度大きな技術革新が起きています。ところが、情報技術に携わる人の中でもその関連分野の専門家しかその技術革新に気が付いていません。それが、情報技術の「見えないイノベーション」です。例えば近年、衆議院議会議録作成システムなどさまざまな実用サービスに音声認

識技術が使われるようになりました。これは、2003年に重み付き有限状態トランスデューサ（WFST: Weighted Finite State Transducer）を音声認識に導入し、200万語彙という超大語彙の音声認識を実現したことが大きな契機となっています⁽¹⁾。こうした情報技術の「見えないイノベーション」をいち早くつかむことが、研究開発においても、サービス開発においても、マーケット開拓においても、成功の鍵であることはいうまでもありません。

果実（み）のなる樹木（き）を育て、世代を重ねて、進化させる

では、こうした「見えないイノベーション」は基礎研究の中から、いつ、どのようにして生まれてくるのでしょうか。基礎研究の営みを図にまとめました。些細な発想の転換や問題の発見が新しい研究の種になります。ふとした気付きの瞬間です。種を播き、水をやるとしばらくして芽が出てくるものがあります。さらに、肥料と陽の光を与えると、いずれ論文や特許として花が咲きます。やがて花は果実となり、摘み取ることができるようになります。この種を播き、樹木（き）を育て果実（み）を摘むまで（図のPhase1）

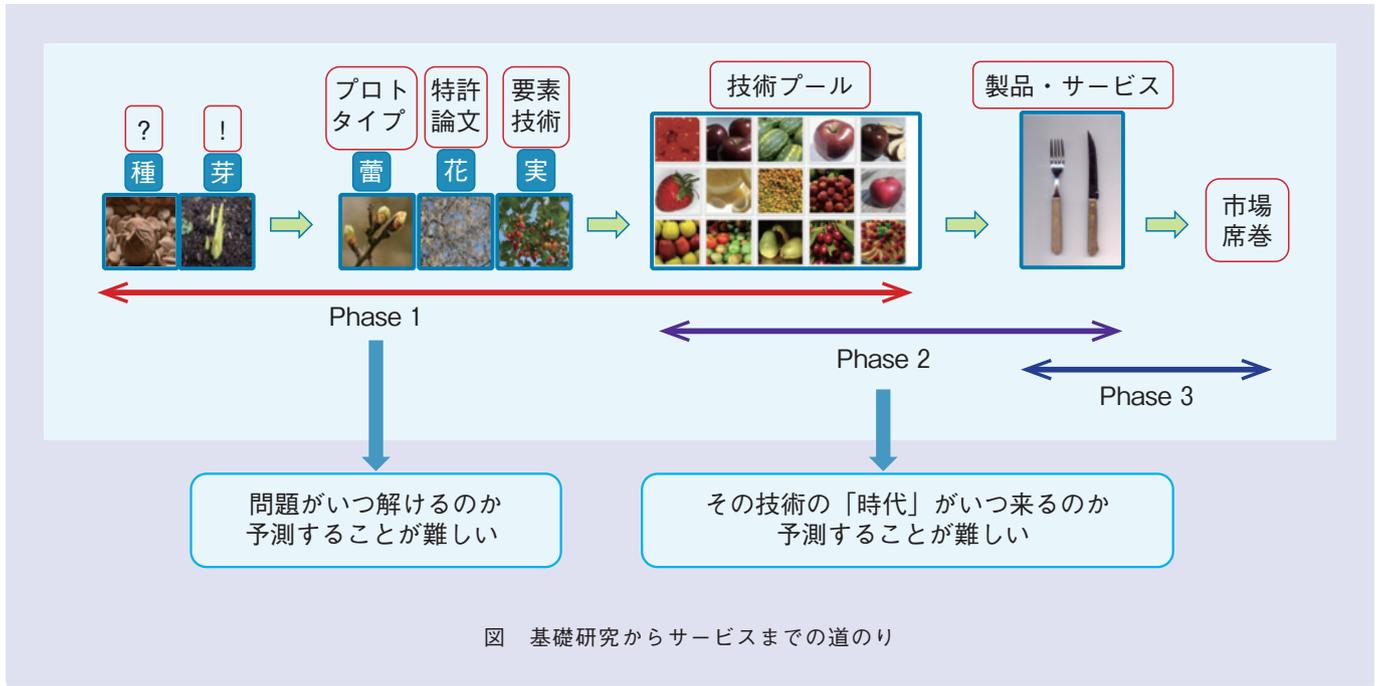


図 基礎研究からサービスまでの道のり

が基礎研究の根幹をなす、もっとも重要な過程です。立派な花が咲いても実をつけなければ意味がありません。また、水や肥料をふんだんにやりさえすれば立派な実がなるというわけでもありません。多様な果実を貯えると同時に、果実から新たな種を採り、再び播くことによって技術のさらなる進化を促します。

基礎研究の使命と価値は、リスクをおそれず困難な課題に挑むことにあります。そのため、その課題がいつ解けるのかを予測することが通常専門家にさえできません。これが、Phase1における最大の問題です。CS研の最新の事例では、人間の感覚特性を利用して牽引力錯覚を引き起こす装置「ぶるなび3」の発明があります。2004年に「ぶるなび1」の開発に成功した時点では、牽引力効果を維持したままぶ

るなびを小型化することは困難だろうと考えていました⁽²⁾。それが、2014年のある日突然解けてしまって、牽引力を損なうことなく、ぶるなび1の20分の1の大きさが実現できたのです。このぶるなび3については、本特集記事『指でつまむと引っ張られる感覚を生み出す装置「ぶるなび3」』で紹介します。

果実が食卓に上り、人々の滋養となる

果実が人々の口に入って初めて滋養となるように、研究成果も技術として使われてこそ、価値を持ちます。CS研の基礎研究成果にも、実社会の中へ入っていく技術が生まれつつあり、その代表例を表にまとめました。これらを分析してみると、種を播いてから10年以上の時間を要しているものが

多いこと、実がなってから食卓に出るまで(図のPhase2)の時間も多様であることが分かります。技術が完成しても時代のニーズに合わなければ使われることはありません。その技術の「時代」がいつ来るのか、それを予想することも多くの場合困難です。これがPhase2における最大の問題です。その「時」が来るまで技術プールを維持し守るとともに、「時」が来たときにすぐに世に出せる体制を整えておかなければなりません。

表に掲げた技術を具体的事例として取り上げてみます。

①ロバストメディア探索(RMS)は、当初、『ウォーリーをさがせ!』という有名な絵本のように、画像の断片を画像全体の中から高速に探し出す画像探索の研究として1993年に始まりました。この技術が、音楽探索、映像探

表 CS研究の成果例

	研究開始	基礎研究の代表的成果	実サービス等への貢献
①	1993年	1995：画像の高速探索（アクティブ探索法） 1998：音・動画の高速探索 2004：雑音・歪みに強いロバストメディア探索 2013：高精度インスタンス探索	2004：携帯電話の音で音楽検索 2007：放送等における音楽著作権処理 2008：ネット投稿動画のコンテンツ特定 2010：スマートフォンによるセカンドスクリーン
②	2000年	2004：超大語彙（200万語）認識デコーダ 2010：音響・言語一体学習モデル	2011：衆議院議会議録システム
③	2003年	2005：基本原理考案 2009：残響制御に基づくサラウンド技術	2009：プロ用残響制御ソフト 2012：民生用サラウンドスピーカ 2013：プロ用サラウンド機器
④	2001年	2003：日本語質問応答システムSAIQA 2009：なぜに答えるNAZEQA	2012：しゃべってコンシェル
⑤	2003年	2009：半教師有り学習による係り受け解析 2010：主辞後置化による事前並べ替え翻訳	2011：ルールベースを超える性能実現 2013：JAPIO中国特許翻訳
⑥	1998年	2007：質感判断の理由解明，Nature誌採録 2012：物体残効錯視の発見	2010：新学術領域「質感情報学」発足
⑦	2003年	2004：牽引デバイスぶるなびの実現（ぶるなび1） 2014：小型化の実現と効果の検証（ぶるなび3）	—

索へと発展し、放送における音楽の著作権管理や投稿動画のコンテンツ特定において主要な役割を果たすようになったのは2008年ごろのことです⁽³⁾。そして現在、さらにインスタンス探索へと発展しつつあり、この詳細については本特集記事『映像中の特定物体を探索するインスタンスサーチ技術』で紹介します。

②WFST型音声認識技術は、前述のWFSTの導入が大きな発展の契機でしたが、近年深層学習というさらに新しい技術が導入され、新たな展開を迎えつつあります⁽⁴⁾。また、③残響制御技術「REVTRINA」は、原理考案から5年以上の歳月を経て、プロ用、民生用のさまざまな機器に導入されつつあります。この技術については、本特集記事『音声をよりクリアに、音楽をより豊かに——残響制御が切り拓く

「音」の世界』で紹介します。

④質問応答技術は、NTTドコモの「しゃべってコンシェル」サービスの基本要素であり、CS研が2001年に研究を開始した質問応答システムSAIQAにそのルーツがあります^{(4)・(5)}。2003年のCS研オープンハウスでは、200万語彙音声認識技術とつなげて音声質問応答のデモンストレーションも実現していました。しかし、この技術が実社会の中で使われるまで、やはり10年の時間を必要としています。この質問応答技術の研究成果は、統計機械学習技術の自然言語処理への導入が2000年ごろから急速に進んだことが大きな契機となっています。これは、⑤機械翻訳技術が従来のルールベース翻訳から統計翻訳へパラダイムシフトする大きな要因にもなりました⁽⁶⁾。

⑥質感情報学は、文部科学省の新学

術領域研究としても2010年に立ち上がり、大きな流れができてつつある学術領域です。これは、2000年に開始したCS研とMITとの共同研究から生まれ、2007年にNatureに採録された成果⁽⁷⁾が契機となっており、現在視覚の質感から五感の質感へと研究が広がつつあります⁽⁸⁾。この詳細については、本特集記事『質感認識の科学と制御——液体質感をもたらす映像中の動き情報の探求』で紹介します。この質感情報学や⑦牽引力錯覚「ぶるなび」は、「時代」が来るのを待っている研究成果であるといえます。このほか、技術として完成途上にある取り組みもたくさんあり、本特集においては、人間の感覚・運動の仕組みについて『身体から潜在的な心を解読するマインドリーディング技術』、量子コンピュータの時代を準備する量子情報処理に関

する取り組みについて『素因数分解ではない量子計算の魅力——量子探索技術の可能性を探る』, 超高臨場感音再生に向けた新しい取り組みについて『光で音をつかまえる——LEDと高速カメラで挑む超多チャンネル音響信号の観測』で, それぞれ紹介します。

基礎研究は「時代」とともに在り

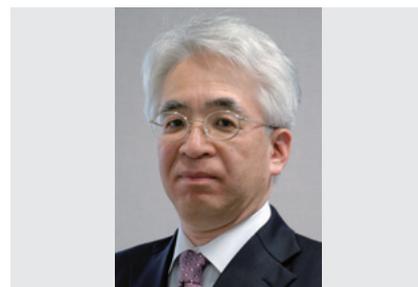
基礎研究の営みは決して世俗から離れた特別なものではありません。時代に寄り添い, 時代とともに存るべきものです。それはかつてベル研究所の幹部であったジョン・ピアースが自身の経験を踏まえ, 「優れた発想や計画はイノベーションには欠かせません。ただ何より時機が肝心です」と述べていることとも符合します⁽⁹⁾。ピアースは, ベル研究所退任後にコンピュータ音楽と呼ばれる学術領域を拓いた人としてもよく知られていますが, 一方, 音声認識研究者の間では1970年前後にベル研究所の音声認識研究を止めた人としても知られています。米国音響学会誌への寄稿⁽¹⁰⁾の中で, “General-purpose speech recognition seems far away. Special-purpose speech recognition is severely limited. It would seem appropriate for people to ask themselves why they are working in the field and what they can expect to accomplish.”と書き, これが米国の音声認識研究に大きな影響を与えました。これは経験豊富な研究マネージャでさえ判断を見誤ることがあるという例であり, 基礎研究運営の難しさを物語っています。実はこのとき, ベル研

究所は音声に関する研究を外国からの客員研究員に任せたのですが, その中の1人が板倉文忠氏であり, そのときの仕事が線スペクトル対 (LSP: Line Spectral Pairs) の研究成果につながるようになるのです。2014年, IEEEマイルストーンに認定されたLSPについては, 本誌「世界に誇れる研究開発成果」の記事『高圧縮音声符号化の必須技術: 線スペクトル対 (LSP)』をご覧ください。

21世紀に入って, 私たちの生活を取り巻く情報環境の変化は激しさを増しており, 図におけるPhase2とPhase3に相当する市場導入はスピードが勝負になっています⁽¹¹⁾。基礎研究においても, その取り組む課題の選択が時代とともにあるのはもちろん, 時代のスピード感に合わせた市場導入への貢献が求められています。NTTが他業種企業と連携したCo-Innovationによる新たな市場創成を目指している中, 基礎研究の価値とその研究成果への期待はむしろ大きくなっているといえるでしょう。基礎研究から生まれる研究成果の1つひとつがイノベーションの貴重な種であり, それらは時代の到来を待っているといえます。

■参考文献

- (1) 久保・小川・堀・中村: “音声と言語の一体型学習に基づく音声認識技術,” NTT技術ジャーナル, Vol.25, No.9, pp.22-25, 2013.
- (2) 雨宮・安藤・何: “五感インタフェースによるノンバーバルコミュニケーション,” NTT技術ジャーナル, Vol.19, No.6, pp.35-37, 2007.
- (3) 柏野・向井・大塚・永野・泉谷・木村・黒住・大和: “高速メディア探索,” NTT技術ジャーナル, Vol.19, No.6, pp.29-32, 2007.
- (4) 前田・磯崎・佐々木・賀沢・平尾・鈴木: “質問応答システム: SAIQA—何でも答える物知り博士,” NTT R&D, Vol.52, No.2, pp.122-133, 2003.
- (5) 東中・貞光・齋藤・小林: “幅広い質問にピンポイントで答える質問応答技術,” NTT技術ジャーナル, Vol.25, No.3, pp.25-28, 2013.
- (6) 永田・須藤・鈴木・秋葉・平尾・塚田: “革新的発展期を迎えた統計翻訳,” NTT技術ジャーナル, Vol.25, No.9, pp.14-17, 2013.
- (7) I.Motoyoshi, S.Nishida, L.Sharan, and E.Adelson: “Image statistics and the perception of surface qualities,” Nature, Vol.447, No.7141, pp.206-209, 2007.
- (8) 渡邊: “触れる感覚の質感・実感に着目したコミュニケーション,” NTT技術ジャーナル, Vol.23, No.9, pp.26-30, 2011.
- (9) ガートナー: “世界の技術を支配する ベル研究所の興亡,” 文藝春秋, 2013.
- (10) J. R. Pierce: “Whither Speech Recognition?,” J. Acoust. Soc. Am., Vol.46, No.2, pp.1049-1051, 1969.
- (11) ボーゲルスタイン: “アップルvs.グーグル: どちらが世界を支配するのか,” 新潮社, 2013.



前田 英作

イノベーションにつながる可能性を秘めた新しい技術がいつ, どこで生まれようとしているのか, それにいち早く気付き, 活用することが, 研究開発においてもサービス創成においても競争を勝ち抜く武器になるはずだ。

◆問い合わせ先

NTTコミュニケーション科学基礎研究所
TEL 0774-93-5000
FAX 0774-93-5015
E-mail maeda.eisaku@lab.ntt.co.jp